

RECEIVED

NOV 20 2002

TECH CENTER 1600/2900



1636 1600

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/636,458

DATE: 11/12/2002

TIME: 17:11:54

Input Set : A:\195675US0.ST25.txt

Output Set: N:\CRF4\11122002\I636458.raw

#8

3 <110> APPLICANT: Matsuzaki, Yumi
 4 Kimura, Eiichiro
 5 Nakamatsu, Tsuyoshi
 6 Kurahashi, Osamu
 7 Kawahara, Yoshio
 8 Sugimoto, Shinichi
 10 <120> TITLE OF INVENTION: Plasmid Autonomously Replicable in Coryneform Bacteria
 12 <130> FILE REFERENCE: 195675US0
 14 <140> CURRENT APPLICATION NUMBER: US 09/636,458
 15 <141> CURRENT FILING DATE: 2000-08-11
 17 <150> PRIOR APPLICATION NUMBER: JP 11-228391
 18 <151> PRIOR FILING DATE: 1999-08-12
 20 <160> NUMBER OF SEQ ID NOS: 20
 22 <170> SOFTWARE: PatentIn version 3.1
 24 <210> SEQ ID NO: 1
 25 <211> LENGTH: 1479
 26 <212> TYPE: DNA
 27 <213> ORGANISM: Corynebacterium thermoaminogenes
 29 <220> FEATURE:
 30 <221> NAME/KEY: CDS
 31 <222> LOCATION: (1)..(1476)
 32 <223> OTHER INFORMATION:

ENTERED

W--> 35 <400> 1
 36 atg act cta gcg gat tcg cca gga aca tac aca gca gat gcg tgg aat 48
 37 Met Thr Leu Ala Asp Ser Pro Gly Thr Tyr Thr Ala Asp Ala Trp Asn
 38 1 5 10 15
 40 tac tcc act gat ctg ttc gac acc cac cct gag ctg gct tta cgc tcc 96
 41 Tyr Ser Thr Asp Leu Phe Asp Thr His Pro Glu Leu Ala Leu Arg Ser
 42 20 25 30
 44 cgg ggt tgg aat cac cag gac gcc gcc gag ttc ctg gcc cac ctg gat 144
 45 Arg Gly Trp Asn His Gln Asp Ala Ala Glu Phe Leu Ala His Leu Asp
 46 35 40 45
 48 cgc agc atg ttt cac ggg tgc ccc acc cgg gat ttc tcc gcg gcc tgg 192
 49 Arg Ser Met Phe His Gly Cys Pro Thr Arg Asp Phe Ser Ala Ala Trp
 50 50 55 60
 52 gtc aaa gac ccg gaa acc gga gaa acc cgc ccc aag ctg cac aga gtt 240
 53 Val Lys Asp Pro Glu Thr Gly Glu Thr Arg Pro Lys Leu His Arg Val
 54 65 70 75 80
 56 ggc acc cgc tca ctt tcc cgg tgc cag tac gtt gcc ctg acc cac ccg 288
 57 Gly Thr Arg Ser Leu Ser Arg Cys Gln Tyr Val Ala Leu Thr His Pro
 58 85 90 95
 60 cag cgc tcc gcg gtg ctg gtc tta gac atc gac atc ccc agc cac cag 336
 61 Gln Arg Ser Ala Val Leu Val Leu Asp Ile Asp Ile Pro Ser His Gln

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/636,458

DATE: 11/12/2002

TIME: 17:11:54

Input Set : A:\195675US0.ST25.txt

Output Set: N:\CRF4\11122002\I636458.raw

62	100	105	110	
64 ggc ggc ggg aac atc gag cac ctt cac ccg cag gtg tac gcc acc ttg				384
65 Ala Gly Gly Asn Ile Glu His Leu His Pro Gln Val Tyr Ala Thr Leu				
66	115	120	125	
68 gag cgt tgg gca cgg gtg gag aaa gcg ccg gcc tgg atc ggg gtg aac				432
69 Glu Arg Trp Ala Arg Val Glu Lys Ala Pro Ala Trp Ile Gly Val Asn				
70	130	135	140	
72 ccg ttg tgc gga aag tgc cag ctc atc tgg tgc att gac ccg gtg ttc				480
73 Pro Leu Ser Gly Lys Cys Gln Leu Ile Trp Cys Ile Asp Pro Val Phe				
74 145	150	155	160	
76 gcc gcc gag ggc acc acc agc tcg aac acc cgc ctg cta gcg gcc acc				528
77 Ala Ala Glu Gly Thr Thr Ser Ser Asn Thr Arg Leu Leu Ala Ala Thr				
78	165	170	175	
80 acc gag gaa atg acc cgt gtg ttc ggc gct gac cag gca ttt tcc cac				576
81 Thr Glu Glu Met Thr Arg Val Phe Gly Ala Asp Gln Ala Phe Ser His				
82	180	185	190	
84 cgg ctg agc cgg tgg ccg ctg cat gtt tct gat gat ccg acc gcg tac				624
85 Arg Leu Ser Arg Trp Pro Leu His Val Ser Asp Asp Pro Thr Ala Tyr				
86	195	200	205	
88 tcc tgg cac tgc cag cac aac cga gtc gat att ctt gat gag ctg atg				672
89 Ser Trp His Cys Gln His Asn Arg Val Asp Ile Leu Asp Glu Leu Met				
90	210	215	220	
92 gag gta gcc cgc acg atg acc gga tca aaa aag ccc aga gag cac gct				720
93 Glu Val Ala Arg Thr Met Thr Gly Ser Lys Pro Arg Glu His Ala				
94 225	230	235	240	
96 cac cag gag ttt tcc agc ggt cgg gca cgg atc gaa gcc gcg cgg aaa				768
97 His Gln Glu Phe Ser Ser Gly Arg Ala Arg Ile Glu Ala Ala Arg Lys				
98	245	250	255	
100 gcc acc gca gag gcc aaa gcg ctt gcc gcc ttg gac gcc acg ctg cct				816
101 Ala Thr Ala Glu Ala Lys Ala Leu Ala Ala Leu Asp Ala Thr Leu Pro				
102	260	265	270	
104 acg gcg ctg gag gca tca ggc gat ctc att gac ggg gtg cgg gtg ttg				864
105 Thr Ala Leu Glu Ala Ser Gly Asp Leu Ile Asp Gly Val Arg Val Leu				
106	275	280	285	
108 tgg gca gca gag ggg cgt gca gcc cgt gat gag aca gcg ttt cgc cat				912
109 Trp Ala Ala Glu Gly Arg Ala Ala Arg Asp Glu Thr Ala Phe Arg His				
110	290	295	300	
112 gcg ttg acc gtg ggt tat cag ctt aaa gcc gca ggt gaa cgc ctg aaa				960
113 Ala Leu Thr Val Gly Tyr Gln Leu Lys Ala Ala Gly Glu Arg Leu Lys				
114 305	310	315	320	
116 gat gcc aag atc att gat gcg tat gag cgt gcc tac aac gtc gcc cag				1008
117 Asp Ala Lys Ile Ile Asp Ala Tyr Glu Arg Ala Tyr Asn Val Ala Gln				
118	325	330	335	
120 gcg gtg gga gct gat ggg cgt gaa ccg gat ctg cct gcc atg cgt gat				1056
121 Ala Val Gly Ala Asp Gly Arg Glu Pro Asp Leu Pro Ala Met Arg Asp				
122	340	345	350	
124 cgt cag acg atg gcc cgc cgt gtg cgc gcc tac gtc gcc aaa ggc cag				1104
125 Arg Gln Thr Met Ala Arg Arg Val Arg Ala Tyr Val Ala Lys Gly Gln				
126	355	360	365	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/636,458

DATE: 11/12/2002

TIME: 17:11:54

Input Set : A:\195675US0.ST25.txt

Output Set: N:\CRF4\11122002\I636458.raw

```

128 ccc acg gtc agc gcc agg agc aca cag acc cag agc agt cgg ggc cgg      1152
129 Pro Thr Val Ser Ala Arg Ser Thr Gln Thr Gln Ser Ser Arg Gly Arg
130      370      375      380
132 aaa gcc ctg gcc acc atg ggc cgc aga ggc ggg caa aaa gcc gct gaa      1200
133 Lys Ala Leu Ala Thr Met Gly Arg Arg Gly Gly Gln Lys Ala Ala Glu
134 385      390      395      400
136 cgc tgg aaa acc gat cct aac ggc aaa tac gcc caa gaa aac cgc caa      1248
137 Arg Trp Lys Thr Asp Pro Asn Gly Lys Tyr Ala Gln Glu Asn Arg Gln
138      405      410      415
140 cga ctc gaa gct gca aac aag cga cgt caa gtc agc tgg aac aaa tac      1296
141 Arg Leu Glu Ala Ala Asn Lys Arg Arg Gln Val Ser Trp Asn Lys Tyr
142      420      425      430
144 gcg agc acg aat tct ggc tac ggt ttc cga cac gta tgg gcc agc ttg      1344
145 Ala Ser Thr Asn Ser Gly Tyr Gly Phe Arg His Val Trp Ala Ser Leu
146      435      440      445
148 gaa aaa tgc cta cgc gac gag caa atc atg gaa gaa aca ggg ctt tca      1392
149 Glu Lys Cys Leu Arg Asp Glu Gln Ile Met Glu Glu Thr Gly Leu Ser
150      450      455      460
152 gaa aaa tgc cta cgc gac gag caa atc atg gaa gaa aca ggg ctt tca      1440
153 Glu Lys Cys Leu Arg Asp Glu Gln Ile Met Glu Glu Thr Gly Leu Ser
154 465      470      475      480
156 tgc caa atc ctt agg ggg gct cac gcc gta gac aga taa      1479
157 Cys Gln Ile Leu Arg Gly Ala His Ala Val Asp Arg
158      485      490
161 <210> SEQ ID NO: 2
162 <211> LENGTH: 492
163 <212> TYPE: PRT
164 <213> ORGANISM: Corynebacterium thermoaminogenes
166 <400> SEQUENCE: 2
168 Met Thr Leu Ala Asp Ser Pro Gly Thr Tyr Thr Ala Asp Ala Trp Asn
169 1      5      10      15
172 Tyr Ser Thr Asp Leu Phe Asp Thr His Pro Glu Leu Ala Leu Arg Ser
173      20      25      30
176 Arg Gly Trp Asn His Gln Asp Ala Ala Glu Phe Leu Ala His Leu Asp
177      35      40      45
180 Arg Ser Met Phe His Gly Cys Pro Thr Arg Asp Phe Ser Ala Ala Trp
181      50      55      60
184 Val Lys Asp Pro Glu Thr Gly Glu Thr Arg Pro Lys Leu His Arg Val
185 65      70      75      80
188 Gly Thr Arg Ser Leu Ser Arg Cys Gln Tyr Val Ala Leu Thr His Pro
189      85      90      95
192 Gln Arg Ser Ala Val Leu Val Leu Asp Ile Asp Ile Pro Ser His Gln
193      100      105      110
196 Ala Gly Gly Asn Ile Glu His Leu His Pro Gln Val Tyr Ala Thr Leu
197      115      120      125
200 Glu Arg Trp Ala Arg Val Glu Lys Ala Pro Ala Trp Ile Gly Val Asn
201      130      135      140
204 Pro Leu Ser Gly Lys Cys Gln Leu Ile Trp Cys Ile Asp Pro Val Phe
205 145      150      155      160

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/636,458

DATE: 11/12/2002

TIME: 17:11:54

Input Set : A:\195675US0.ST25.txt

Output Set: N:\CRF4\11122002\I636458.raw

```

208 Ala Ala Glu Gly Thr Thr Ser Ser Asn Thr Arg Leu Leu Ala Ala Thr
209                               165                               170                               175
212 Thr Glu Glu Met Thr Arg Val Phe Gly Ala Asp Gln Ala Phe Ser His
213                               180                               185                               190
216 Arg Leu Ser Arg Trp Pro Leu His Val Ser Asp Asp Pro Thr Ala Tyr
217                               195                               200                               205
220 Ser Trp His Cys Gln His Asn Arg Val Asp Ile Leu Asp Glu Leu Met
221                               210                               215                               220
224 Glu Val Ala Arg Thr Met Thr Gly Ser Lys Lys Pro Arg Glu His Ala
225 225                               230                               235                               240
228 His Gln Glu Phe Ser Ser Gly Arg Ala Arg Ile Glu Ala Ala Arg Lys
229                               245                               250                               255
232 Ala Thr Ala Glu Ala Lys Ala Leu Ala Ala Leu Asp Ala Thr Leu Pro
233                               260                               265                               270
236 Thr Ala Leu Glu Ala Ser Gly Asp Leu Ile Asp Gly Val Arg Val Leu
237                               275                               280                               285
240 Trp Ala Ala Glu Gly Arg Ala Ala Arg Asp Glu Thr Ala Phe Arg His
241                               290                               295                               300
244 Ala Leu Thr Val Gly Tyr Gln Leu Lys Ala Ala Gly Glu Arg Leu Lys
245 305                               310                               315                               320
248 Asp Ala Lys Ile Ile Asp Ala Tyr Glu Arg Ala Tyr Asn Val Ala Gln
249                               325                               330                               335
252 Ala Val Gly Ala Asp Gly Arg Glu Pro Asp Leu Pro Ala Met Arg Asp
253                               340                               345                               350
256 Arg Gln Thr Met Ala Arg Arg Val Arg Ala Tyr Val Ala Lys Gly Gln
257                               355                               360                               365
260 Pro Thr Val Ser Ala Arg Ser Thr Gln Thr Gln Ser Ser Arg Gly Arg
261                               370                               375                               380
264 Lys Ala Leu Ala Thr Met Gly Arg Arg Gly Gly Gln Lys Ala Ala Glu
265 385                               390                               395                               400
268 Arg Trp Lys Thr Asp Pro Asn Gly Lys Tyr Ala Gln Glu Asn Arg Gln
269                               405                               410                               415
272 Arg Leu Glu Ala Ala Asn Lys Arg Arg Gln Val Ser Trp Asn Lys Tyr
273                               420                               425                               430
276 Ala Ser Thr Asn Ser Gly Tyr Gly Phe Arg His Val Trp Ala Ser Leu
277                               435                               440                               445
280 Glu Lys Cys Leu Arg Asp Glu Gln Ile Met Glu Glu Thr Gly Leu Ser
281                               450                               455                               460
284 Glu Lys Cys Leu Arg Asp Glu Gln Ile Met Glu Glu Thr Gly Leu Ser
285 465                               470                               475                               480
288 Cys Gln Ile Leu Arg Gly Ala His Ala Val Asp Arg
289                               485                               490
292 <210> SEQ ID NO: 3
293 <211> LENGTH: 1479
294 <212> TYPE: DNA
295 <213> ORGANISM: Corynebacterium thermoaminogenes
297 <220> FEATURE:
298 <221> NAME/KEY: CDS
299 <222> LOCATION: (1)..(1476)

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/636,458

DATE: 11/12/2002

TIME: 17:11:54

Input Set : A:\195675US0.ST25.txt

Output Set: N:\CRF4\11122002\I636458.raw

300 <223> OTHER INFORMATION:

W--> 303 <400> 3

304	atg act cta gcg gat tgc cca gga aca tac aca gca gat gcg tgg aat	48
305	Met Thr Leu Ala Asp Ser Pro Gly Thr Tyr Thr Ala Asp Ala Trp Asn	
306	1 5 10 15	
308	tac tcc act gat ctg ttc gac acc cac cct gag ctg gct tta cgc tcc	96
309	Tyr Ser Thr Asp Leu Phe Asp Thr His Pro Glu Leu Ala Leu Arg Ser	
310	20 25 30	
312	cgg ggt tgg aat cac cag gac gcc gca gag ttc ctg gcc cac ctg gat	144
313	Arg Gly Trp Asn His Gln Asp Ala Ala Glu Phe Leu Ala His Leu Asp	
314	35 40 45	
316	cgc agc atg ttt cac ggg tgc ccc acc cgg gat ttc tcc gcg gcc tgg	192
317	Arg Ser Met Phe His Gly Cys Pro Thr Arg Asp Phe Ser Ala Ala Trp	
318	50 55 60	
320	gtc aaa gac ccg gaa acc gga gaa acc cgc ccc aag ctg cac aga gtt	240
321	Val Lys Asp Pro Glu Thr Gly Glu Thr Arg Pro Lys Leu His Arg Val	
322	65 70 75 80	
324	ggc acc cgc tca ctt tcc cgg tgc cag tac gtt gcc ctg acc cac ccg	288
325	Gly Thr Arg Ser Leu Ser Arg Cys Gln Tyr Val Ala Leu Thr His Pro	
326	85 90 95	
328	cag cgc tcc gcg gtg ctg gtc tta gac atc gac atc ccc agc cac cag	336
329	Gln Arg Ser Ala Val Leu Val Leu Asp Ile Asp Ile Pro Ser His Gln	
330	100 105 110	
332	gcc ggc ggg aac atc gag cac ctt cac ccg cag gtg tac gcc acc ttg	384
333	Ala Gly Gly Asn Ile Glu His Leu His Pro Gln Val Tyr Ala Thr Leu	
334	115 120 125	
336	gag cgt tgg gca cgg gtg gag aaa gcg ccg gcc tgg atc ggg gtg aac	432
337	Glu Arg Trp Ala Arg Val Glu Lys Ala Pro Ala Trp Ile Gly Val Asn	
338	130 135 140	
340	ccg ttg tgc gga aag tgc cag ctc atc tgg tgc att gac ccg gtg ttc	480
341	Pro Leu Ser Gly Lys Cys Gln Leu Ile Trp Cys Ile Asp Pro Val Phe	
342	145 150 155 160	
344	gcc gcc gag ggc acc acc agc tgc aac acc cgc ctg cta gcg gcc acc	528
345	Ala Ala Glu Gly Thr Thr Ser Ser Asn Thr Arg Leu Leu Ala Ala Thr	
346	165 170 175	
348	acc gag gaa atg acc cgt gtg ttc ggc gct gac cag gca ttt tcc cac	576
349	Thr Glu Glu Met Thr Arg Val Phe Gly Ala Asp Gln Ala Phe Ser His	
350	180 185 190	
352	cgg ctg agc cgg tgg ccg ctg cat gtt ttt gat gat ccg acc gcg tac	624
353	Arg Leu Ser Arg Trp Pro Leu His Val Phe Asp Asp Pro Thr Ala Tyr	
354	195 200 205	
356	tcc tgg cac tgc cag cac aac cga gtc gat att ctt gat gag ctg atg	672
357	Ser Trp His Cys Gln His Asn Arg Val Asp Ile Leu Asp Glu Leu Met	
358	210 215 220	
360	gag gta gcc cgc acg atg acc gga tca aaa aag ccg aga aag cac gct	720
361	Glu Val Ala Arg Thr Met Thr Gly Ser Lys Lys Pro Arg Lys His Ala	
362	225 230 235 240	
364	cac cag gag ttt tcc agc ggt cgg gca cgg atc gaa gcc gcg cgg aaa	768
365	His Gln Glu Phe Ser Ser Gly Arg Ala Arg Ile Glu Ala Ala Arg Lys	

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/636,458

DATE: 11/12/2002

TIME: 17:11:55

Input Set : A:\195675US0.ST25.txt

Output Set: N:\CRF4\11122002\I636458.raw

L:35 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:32
L:303 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:3,Line#:300
L:571 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:5,Line#:568
L:839 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:7,Line#:836
L:1091 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:9,Line#:1088